Passive Night Vision Sensor Comparison for Unmanned Ground Vehicle Stereo Vision Navigation

Ken Owens and Larry Matthies
Jet Propulsion Laboratory, 4800 Oak Grove Dr., Pasadena CA, 91109
kowens@robotics.jpl.nasa.gov lhm@robotics.jpl.nasa.gov

Abstract

One goal of the "Demo III" unmanned ground vehicle program is to enable autonomous nighttime navigation at speeds of up to 10 m.p.h. To perform obstacle detection at night with stereo vision will require night vision cameras that produce adequate image quality for the driving speeds, vehicle dynamics, obstacle sizes, and scene conditions that will be encountered. This paper analyzes the suitability of four classes of night vision cameras (3-5 μm cooled FLIR, 8-12 μm cooled FLIR, 8-12 μm uncooled FLIR, and image intensifiers) for night stereo vision, using criteria based on stereo matching quality, image signal to noise ratio, motion blur, and synchronization capability. We find that only cooled FLIRs will enable stereo vision performance that meets the goals of the Demo III program for nighttime autonomous mobility.

1.0 Introduction

Experimental Unmanned Vehicles (XUV) of the 2500 lb weight class are being developed under the Demo III program to significantly broaden the operating range of unmanned ground vehicles to include night operations in cross country environments. During this program, the XUVs are to autonomously drive during the night at speeds of up to 10 m.p.h. while using stereo vision to detect and avoid obstacles such as 12 inch high bumps and 24 inch wide ditches.

The basic feasibility of doing nighttime stereo vision was demonstrated during the Demo II program, see [1] and [2]. A legacy pair of night vision Indium Antimonide (InSb) forward looking infrared (FLIR) cameras was installed on the Jet Propulsion Laboratory (JPL) High Mobility Multipurpose Wheeled Vehicle (HMMWV) to perform further feasibility analysis of 24 hour round-the-clock stereo vision based obstacle detection. The images in figure 1 were taken at 6:00 am, just before dawn, with the JPL night vision stereo system at Aberdeen Proving Ground; they show an 11 inch rock on a dirt road 37 ft. in front of the night vision cameras with estimated range and height maps clearly showing this same rock.

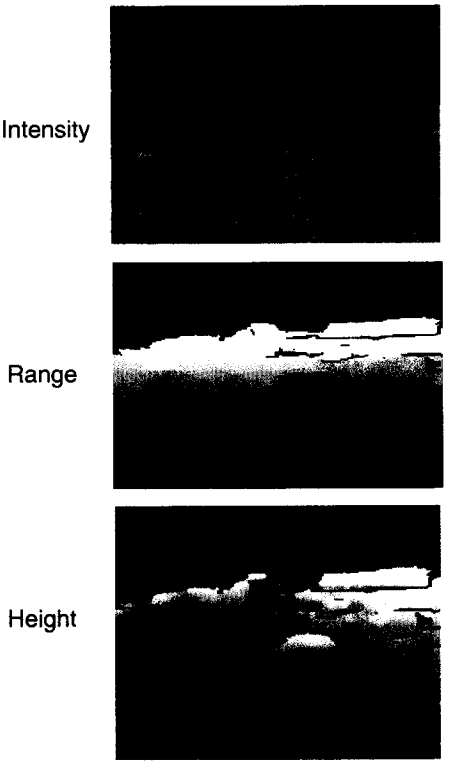


Figure 1: 6:00 am InSb FLIR image of a rock, 37 ft. from the camera with corresponding stereo range and height estimates clearly showing the same rock.

The XUVs in the Demo III program are to provide an autonomous forward observing capability for ground troops, during which they travel ahead reporting on remote battlefield conditions. In this arena, non-emissive range sensors with a 24 hour operating capability are paramount. To achieve this capability, CCD and FLIR cameras will be used. The use of CCD cameras for daytime stereo vision is a mature science; however, much work needs to be done to understand which night vision cameras are suitable for nighttime XUV stereo vision.

Commercially available night imaging technology falls

into four main categories. Three of these technology groups use focal planes that operate in the 3-5 and 8-12 micron band transmission windows. The 3-5 micron band focal planes are available in cooled photon counting formats and are most often composed of Indium Antimonide (InSb) and Platinum Silicide (PtSi). Cooled 8-12 micron band focal planes also are photon counting devices and are readily available in Gallium Arsenide (GaAs) or Mercury Cadmium Teluride (MCT). Uncooled 8-12 micron detectors, such as microbolometers and pyroelectric devices are available; however instead of counting photons, they are physically sensitive to infrared radiation using technology such as arrays of infrared sensitive resistors. In addition to these infrared imaging devices there are also image intensifiers that work during the night in the visible spectrum by using photomultiplier devices.

Previous work in [3] and [4] conducted at the Night Vision and Electronics Sensors Directorate (NVESD) has quantified the performance of night vision technology at the sensor level. However, these studies alone do not provide the necessary guidance for selecting the best night vision technology for XUV stereo vision. In this paper, we address the issues necessary for determining which of the existing commercial off the shelf (COTS) night vision technologies are suitable for the Demo III nighttime stereo application. To make this determination, we compared images from the above four basic night vision sensor groups with a series of derived camera performance requirements. The derived requirements specify the allowable stereo disparity error, signal to noise ratio, motion blur limit, and exposure timing necessary for successful XUV night time stereo system operation. This paper will be divided into four sections, addressing these issues.

2.0 Stereo Disparity Error Analysis

To investigate image quality requirements for night vision stereo range sensors on board XUVs, a series of images was collected with representative cameras from the four major groups of COTS night vision technologies: (1) 3-5 micron cooled Indium Antimonide (InSb), (2) 8-12 micron cooled Mercury Cadmium Telluride (MCT) and 8-12 micron cooled GaAs Quantum Well Infrared Photodetectors (QWIP), (3) 8-12 micron uncooled microbolometers and pyroelectric detectors and (4) third generation image intensifiers. Pairs of images from these cameras were transformed into false stereo pairs by shifting one of the images and the standard JPL stereo algorithm [5] was run, see figure 3.

When stereo matching is performed on this type of image pair the range to each point in the "scene" is constant since the disparity for each pair of image points is the same. This analysis allows us to estimate the disparity standard deviation using the difference of the actual and measured disparities. In figure 3 we show one of the intensity images from a false stereo pair and the corresponding disparity estimates from each of night vision sensors tested. Since the actual disparity is constant, variations in the disparity image indicate poor stereo performance and in fact the quality of the stereo matching is inversely proportional to the variation in the disparity image. The average color of the disparity images represents the constant disparity being estimated while black regions indicate where stereo matching failed. The disparity error statistics for all of the tested night vision cameras are shown in the table 1. These disparity estimates were computed from 480X512 digitized images using only level 1 resolution (level 0 represents full digitizer resolution, level 1 half resolution, etc.) with a 7X7 correlation window, left-right line of sight error checking and region filtering; disparity outliers were eliminated during the estimation process.

To understand the implications of the empirical results in table 1, image quality standards must be derived from the XUV mobility requirements. To begin this derivation, consider the problem of detecting a far away ditch from near ground level where one pixel observes the ground in front of the ditch and the next higher pixel in the column observes the far side of the ditch.

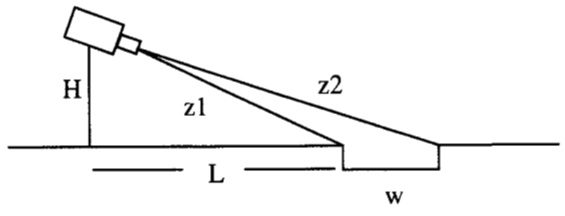
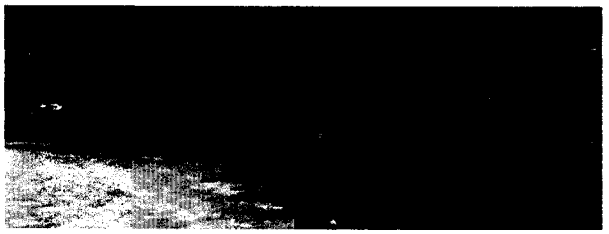


Figure 2: Camera geometry.

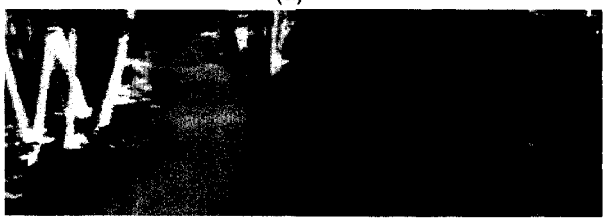
In this situation, the camera height is much less than the range, so the ditch width is approximately equal to the difference in the actual range values shown in figure 2. For good detection, the difference in the measured range values should be some appreciable fraction of the actual ditch width. Put mathematically, the standard deviation in the range Z must be small enough to make the following probability large

$$P(Z_2 - Z_1 > t \,|\, z_2 - z_1 = w) > 0.97$$

where, Z represents random range measurements, t is a fraction of the ditch width, z represents the actual range values, and w represents the width of the ditch.



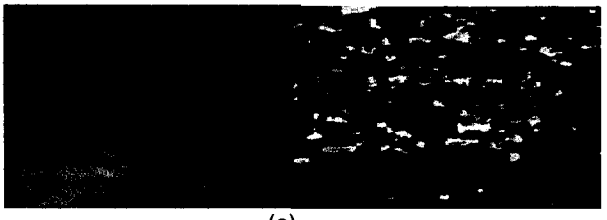
(a)



(b)

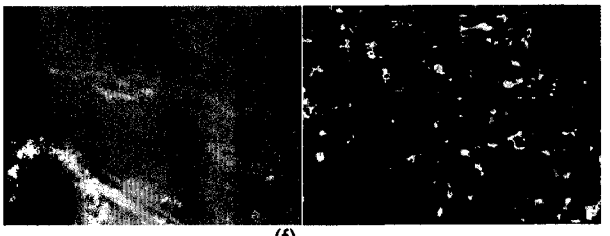
(c)

(d)



(e)

Figure 3: Images and disparity from false stereo pairs for a) 3-5 cooled Galileo InSb FLIR, b) 3-5 cooled Radiance I InSb FLIR, c) 3-5 cooled InSb Infrcam FLIR, d) 8-12 cooled MCT MilCam FLIR. e) 8-12 cooled GaAs QWIP.



(f)

(g)

(h)

(i)

Figure 3(continued): Images and disparity from false stereo pairs for f) 8-12 uncooled Bolometer image captured during the day, g) 8-12 uncooled TI BST Pyroelectric nighttime image h) and i) XR/M Stanford Photonics Gen III Ultra image intensifier images in a nighttime urban environment and on a very dark night in the wilderness. The average color in the disparity maps represents the constant disparity value being estimated; variations about this color indicate disparity estimation errors and black regions indicate the absence of disparity estimates.

Night Vision Camera	Disparity Standard Deviation (Pixels)	Percent Stereo Matches
Cooled		
3-5 InSb, Amber/Raytheon Galileo	0.08	91.6
3-5 InSb, Amber/Raytheon Radiance I	0.06	99.9
3-5 InSb, Inframetrics InfraCam***	0.12	99.5
8-12 MCT, Inframetrics MilCam***	0.21	96.2
8-12 GaAs, Inframetrics /JPL QWIP	0.22	92.1
Uncooled		
8-12 Bolometer, Amber Sentinel *	0.21	90.4
8-12 BST, TI Driver Vision Enhancer	0.19	78.0
Intensifier		
Stanford Photonics XR/M Gen III, Intensifier	2.20	13.0
Stanford Photonics XR/M Gen III, Intensifier **	9.55	1.0

Table 1: Night vision sensor disparity errors for detectors composed of Indium Antimonide (InSb), Mercury Cadmium Telluride (MCT), Gallium Arsenide (GaAs) Quantum Well Infrared Photodetector (QWIP), Microbolometer, Berrium Strontium Titanate (BST) Pyroelectric, and Gen III image intensifier. *Daytime, **dark night, and ***NVESD image acquisition.

By subtracting the actual range measurements and dividing by the standard deviation in Z, this inequality can be transformed into

$$P\left(\frac{(Z_2-z_2)}{\sigma_z}-\frac{(Z_1-z_1)}{\sigma_z}>\frac{(t-\Delta z)}{\sigma_z}\middle|\Delta z=w\right)>0.97$$

It has been shown, that to a good approximation, the range measurements Z are Gaussian [5]. Thus, the quantity on the left hand side of the above inequality is normally distributed with a variance of 2 and we can write,

$$P\left(\Phi>\frac{(t-\Delta z)}{\sqrt{2}\sigma_z}\middle|\Delta z=w\right)>0.97\Rightarrow\frac{(t-w)}{\sqrt{2}\sigma_z}<-2$$

where Φ is the standard normal. With a threshold of $t=3w/4$,

$$\sigma_z<\frac{w}{8\sqrt{2}}$$

which guarantees a per pixel probability of detection that is greater than 97 percent. For a stereo system with baseline, b, focal length, f, and a given disparity standard deviation, it has been shown [5]

$$\sigma_Z=\frac{\sigma_d}{bf}Z^2$$

where Z is the range. The last two relationships imply that, the ratio of the disparity standard deviation and the focal length is given by

$$\frac{\sigma_d}{f}<\frac{wb}{8\sqrt{2}Z^2}$$

Inserting the Demo III night driving parameters of, w=0.6 meters, b=0.35 meters, Z=5.5 meters (The Z value use here is the stopping distance. See section 4.0), f=16 mm and a pixel size of 30 microns, we find the disparity standard deviation limit of

$$\sigma_d<0.3 pixels$$

There is the possibility of using a smaller threshold and relaxing this disparity requirement. However, the per pixel false alarm rate associated with the above detection threshold is about one false alarm in five minutes, assuming a 10 Hz stereo machine processing 256X256 pixels. Thus we would not want to use a smaller threshold. Other avenues for relaxing this disparity standard deviation requirement such as increasing the focal length or the baseline are also limited because the focal length is restricted by the mobility field of view requirements and the baseline is restricted by the physical size of the pan-tilt head. The size of pan-tilt would have to at least increase by a factor of 7 to accommodate the poor disparity estimation performance of image intensifiers, effectively increasing the pan moment of inertia beyond acceptable limits.

3.0 Sensor Signal to Noise Analysis

In this section both empirical and theoretical tech-

niques will be used to estimate the signal to noise performance of COTS night vision sensors. We begin with an empirical analysis of images collected with the night vision sensors described in the previous section and follow with theoretical estimates of photon signal to noise ratios, and camera signal to noise ratios as a function of scene temperature difference.

Camera	Mean Intensity	Intensity Noise Level	Signal/ Noise
Cooled			
3-5 InSb, Galileo	92.7	2.2	41.3/1
3-5 InSb, Radiance	122.1	3.5	34.4/1
3-5 InSb, InfraCam***	43.1	0.91	47.3/1
8-12 MCT, MilCam***	41.8	0.88	47.2/1
8-12 GlAs, QWIP	83.0	1.4	58.1/1
Uncooled			
8-12, Bolometer*	142.9	1.95	73.2/1
8-12 BST, DVE	182.7	11.3	16.2/1
Intensifier			
Gen III, Intensifier**	50.9	10.6	4.8/1
Gen III, Intensifier	97.0	14.4	6.7/1

Table 2: Night vision camera experimental signal to noise ratio. *Daylight, ** very dark night, and ***NVESD image capture.

The images collected from the COTS night vision sensors provide an empirical method of evaluating the signal to noise performance of a given technology. By differencing a pair of images from a given sensor, the intensity standard deviation can be estimated. Taking the ratio of the average intensity and this standard deviation yields an estimate of the signal to noise ratio.

Table 2 summarizes the results of such an analysis of the COTS night vision images. To graphically illustrate the effect of the signal to noise ratio on stereo performance, we have plotted the percentage of stereo matches vs. the signal to noise ratio in figure 4. Note that daytime FLIR data from an uncooled 8-12 micron bolometer produced the highest point on the graph in figure 4. This high signal to noise ratio is expected late in day when this image was taken and does not indicate that this uncooled camera has better performance than cooled night vision sensors. Figure 4 indicates that in order to have an 90 percent stereo match level, night vision devices must have a signal to noise ratio of at least 30 to 1.

Many detectors in use for night vision count photon arrivals over a rectangular grid of pixels. Devices such as image intensifiers amplify and count photon arrivals in the visible spectrum while long and short wave infrared detectors count infrared photon arrivals. The inherent noise in all signals composed of photons can be modeled as a Poisson process [4], which gives the probability of getting a particular number of photons in a given time interval. This distribution has a mean number of photons arrivals equal to the average arrival rate and a variance which also equals this rate (a special property of Poisson processes). This implies that the signal to noise ratio for a Poisson process is given by the square root of the arrival rate.

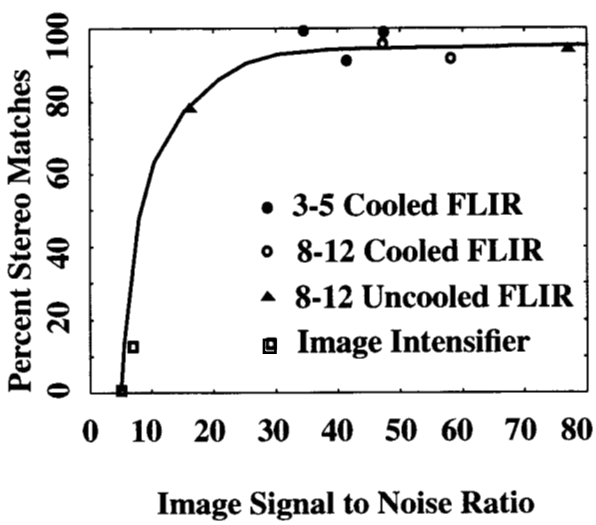


Figure 4: Empirical probability of making a stereo match as a function of signal to noise ratio.

The Poisson process photon arrival rate per pixel, R, is estimated by the ratio of the image irradiance and the photon energy. In the case of normal incidence, accord-

ing to [7], the photon arrival rate is given by

$$R = L\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \frac{A}{E_p}$$

where L is the scene radiance, f/d is the f-number of the optics, A is the pixel area and E_p is the photon energy. The scene radiance will be approximated by making a Lambertian surface assumption in both the visible and infrared cases. In the visible case, the scene radiance is

$$L = \frac{E_0 \rho}{\pi}$$

where E_0 is the scene irradiance [7] in watts/meter^2 and ρ is the albedo, for which we use the typical value [6] of 0.3. In the infrared case the scene radiance is given by

$$L = \frac{E_0 \varepsilon}{\pi}$$

where ε is the emissivity, which for natural terrain is approximately 0.90 [8]. Finally, the signal to noise ratio is

$$\frac{S}{N} = \sqrt{R}$$

Using this expression and scene radiances from [9], we can generate table 3 showing signal to noise ratios which suggest that during low light conditions, photon arrival noise will dominate the signal of image intensifiers. Thus, no matter how good a photomultiplier is used, for short exposure times on dark nights, image intensifiers will produce low quality images making them unacceptable as XUV nighttime stereo vision sensors for the Demo III program; figure 5 shows simultaneously captured cooled FLIR, uncooled FLIR, and Gen III Ultra image intensifier images collected in the wooded canyon area behind JPL. It was a partially cloudy 50 F night where the half moon that evening was not shining in the canyon due to the clouds and geometry. The image intensifier was set with a 16 millisec integration time yielding a predicted signal to noise ratio of approximately 3/1, in good agreement with the measured value of 4.8/1. The cooled FLIR was set to an integration time of 0.25 millisec, implying a signal to noise ratio of 464/1, more than an order of magnitude larger than the measured value, indicating that other noise sources are dominant. However, at temperatures likely to be encountered in cross country environments, the combined infrared signal noise is still within reasonable limits.

The combined infrared sensor noise is characterized by the noise equivalent temperature difference (NEDT). The NEDT for a typical cooled infrared camera is 0.025K at 23 C whereas the NEDT for an uncooled system is on the order of 0.1K at 23 C. The NEDT is the smallest temperature difference that a given camera can resolve. Theoretical expressions [4] suggest that the NEDT increases with decreasing temperature implying a decrease in camera sensitivity. However, in the following argument, we assume that the NEDTs are constant when establishing the signal to noise ratios as a function of scene temperature difference, thereby slightly over estimating camera sensitivity.

As night progresses, temperature differences tend to equalize, compressing the dynamic range of the night vision device into smaller temperature ranges. Since the noise is constant or increases, it occupies a larger portion of the dynamic range thereby reducing the signal to noise ratio. For example, some night vision cameras output images as 8 bit pixels, i.e. with 256 grey levels. A person at night might have a 20 C temperature difference from the background. If this represents full scale in the image, there are 12.8 grey levels per degree. Thus, the NEDT of a cooled infrared camera would be represented by 0.32 of a grey level while the NEDT for an uncooled camera would be represented by 1.28 grey levels. Since these noise levels are represented by only about one grey scale value, both cameras are operating at near optimal levels for 8 bit images.

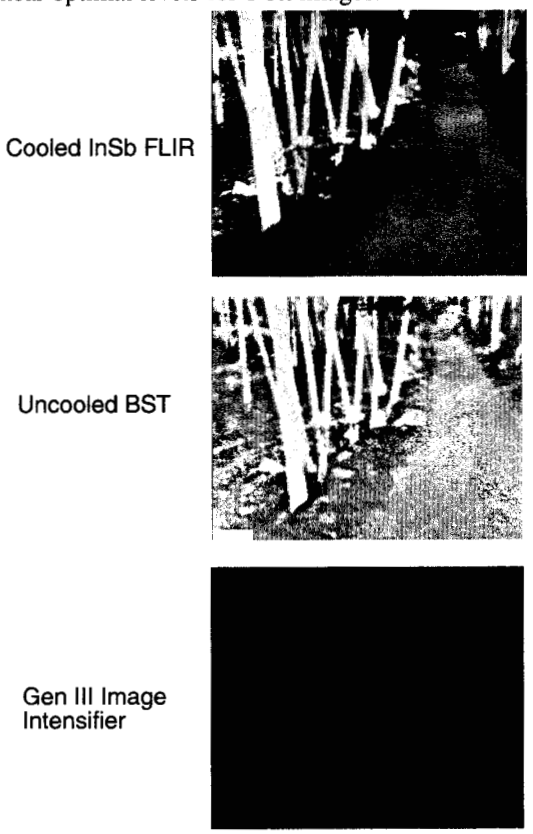


Figure 5: Simultaneously captured nighttime images by an InSb cooled FLIR, BST uncooled FLIR and a Gen III Ultra image intensifier.

Condition/ Spectral Region	Radiance (W/m^2/ sr)	S/N @ 1 ms	S/N @ 16 ms	S/N @ 33 ms
1/4 Hr Past Sunset / Visible	4.4E-4	12/1	49/1	70/1
Clear Full Moon/ Visible	1.4E-5	2/1	9/1	12/1
Clear 1/4 Moon / Visible	1.4E-6	0.7/1	3/1	4/1
Clear Moonless/ Visible	4.7E-7	0.2/1	0.9/1	1/1
0 C/ 3-5 micron FLIR	2.0E-1	877/1	N/A	N/A

Table 3: Estimated photon arrival signal to noise ratios for image intensifiers under various illumination conditions and for a 3-5 micron FLIR at 0 C. An albedo of 0.3 and an emissivity of 0.9 were used in computing the scene radiance and the number of photons per second per pixel were computed with f-number 1.8, and a pixel size of 13 microns for the image intensifier and f-number 2.3 and a pixel size of 38 microns for the FLIR. The last two columns for the FLIR are unanswered since these integration times would saturate the detector.

Table 4 summarizes calculations like the example above for different temperature differences. We see from this table that temperature differences on the order of a degree produce signal to noise ratios of 40/1 in cooled and 12/1 in uncooled sensors. In the previous section, figure 4 showed that a 30/1 signal to noise ratio corresponds to approximately a 90 percent stereo match level and that the percentage of stereo matches decreases sharply to unacceptable levels for signal to noise ratios below 30/1.

To estimate how often actual temperature differences are near 1 C in cross country environments, we measured outdoor temperature differences on the Aberdeen Proving Grounds at various times of the day and night. This data was collected as part of a XUV 24 hour round-the-clock stereo vision and LADAR performance evaluation [10]; figure 1 shows a typical image from this data collection. In table 5, we see that the temperature differences between a wide range of materials have temperature differences near 1 C for long periods of time near dawn and dusk. The period near dawn is characterized by a change in thermal contrast during which temperature differences are low. As scene temperatures ranges become compressed into the one degree range during low thermal contrast periods, cooled night vision sensors with signal to noise ratios on the order of 40/1 will produce acceptable stereo performance as indicated by figure 4. However, for the same scene temperature range, uncooled night vision sensors will have only a 12/1 signal to noise ratio, producing unacceptable stereo range data. However, during high temperature contrast periods where the scene temperature range is greater than 3 C, uncooled night vision sensors should produce adequate stereo range data while the camera is at rest, as in the daylight FLIR image shown in figure 3f. In the next section we will see that both low sensitivity and

Scene Temperature Difference	Night Vision Camera Noise Level (out of 256 grey levels)	Signal to Noise Ratio
Cooled Cameras		
20 C	0.25	1000/1
5 C	1	256/1
3 C	2	128/1
1 C	6	40/1
0.5 C	12	21/1
Uncooled Cameras		
20 C	1	256/1
5 C	4	64/1
3 C	7	36/1
1 C	20	12/1
0.5 C	40	6/1

Table 4: Estimated image signal to noise ratio vs. scene temperature difference.

Time	Rock Face vs. Road	Rock Top vs. Road	Grass vs. Road	Trees vs. Road
5:55am	-0.3C	1.0C		
6:20am	0.1C	-0.1C		-0.8C
6:50am		0.4C	0.7C	1.1C
8:15am		-0.3C	1.9C	4.7C
5:17pm	1.2C	-2.2C		-1.1C
9:10pm	0.3C	0.5C		1.6C

Table 5: Aberdeen Proving Ground measured temperature differences.

long exposure times make uncooled night vision sensors unsuitable for nighttime XUV stereo vision on the move.

4.0 Motion Blur Analysis

There is a basic trade-off in any motion blur analysis between system complexity and hardware. If one is willing to invest a considerable amount of time and money, a system could be designed that would eliminate most of the vehicle induced motion blur. This system would mount the camera on a motion stabilized platform and used active control of image pointing on time scales shorter than the exposure time to reduce motion blur. On the other hand, the camera could be mounted directly to a non-stabilized pan-tilt mechanism and the effects of vehicle motion on image quality would be compensated for by using a fast exposure time. While the first approach is an interesting option, the second approach is more practical in the short term since it is realizable with existing camera hardware for cross country driving speeds. In this section we derive requirements on the image exposure time necessary for successful nighttime XUV stereo vision based on this second method.

The first issue in making a blur analysis is to establish the instantaneous field of view (IFOV) required for the system. The IFOV is the angle subtended by a pixel as viewed from the optical center. This angle is in turn determined by the size and distance of obstacles that must be resolved and how many pixels are required to fall on an obstacle for reliable detection. For cross country navigation, ditches or so called negative obstacles that are most difficult to detect. Thus let us begin by deriving the necessary IFOV for a night vision system to detect a ditch of a given width at a given distance (see figure2). For small angles, the result is

$$IFOV = \frac{HW}{NL^2}$$

where H is the sensor height, W is the width of the negative obstacle, N is the number of pixels on target, and L is the distance from the vehicle to the obstacle. The factor HW/L is the effective size of a negative obstacle of width W as seen from a height H and a distance L. This follows from the fact that the apparent size of a negative obstacle is the width W multiplied by the sine of the angle below the horizon of the pixel line of sight, which for small angles is approximately H/L. Again using small angles, the angle subtended by the obstacle, N*IFOV, is equal to the ratio of the effective obstacle size and distance to the obstacle, from which the formula follows. For the Demo III parameters of H=1.0m, W=0.6m, N=10pixels, L=5.5m (see the stopping distance formula later in this section) we get IFOV=1.98 millirad. The proposed Demo III system with a 16mm lens and 30 μm pixels has a 1.88 millirad resolution, just meeting this requirement.

Now as the camera moves through space, the image moves across the focal plane and this motion causes image blur. There are many contributors to this image blur such a vehicle pitching, pan-tilt motion, and as mentioned vehicle translation. Let us begin with a discussion of vehicle translation. As the vehicle travels at velocity, U, the image points on the focal plane move at an angular rate of

$$\omega = \frac{HU}{L^2}$$

For small angles, the quantity HU/L is equal to the component of the vehicle velocity perpendicular to the direction of observation. The distance to the observed point, L, multiplied by the angular rate of motion must equal this velocity, and the formula follows.

As a criteria for motion blur, we will say that an image has blurred if there has been motion on the focal plane equivalent to 1/2 of the IFOV in an exposure time, τ. Thus, we have the motion blur restriction

$$\omega\tau < \frac{1}{2} IFOV$$

When considering only translation, and substituting the above expressions, this equation becomes

$$\tau < \frac{\frac{HW}{NL^2}}{2\frac{HU}{L^2}} = \frac{W}{2NU}$$

For a 2ft ditch seen while traveling at 5 m/s (about 10

m.p.h.) with 10 pixels on the obstacle,

$$\tau < 6ms$$

Since the fastest uncooled night vision systems currently operate with 16 millisec exposure times, this requirement makes them unsuitable for stereo matching while directly mounted on moving XUVs. However, many cooled systems have exposure time on the order of one millisecond, well below this threshold.

As mentioned there are other contributors to motion blur such as vehicle pitch and pan-tilt motion. Let us model the pitch angular velocity

$$\frac{U\sin(C_1)}{D}$$

and the pan-tilt angular velocity by

$$\frac{U\sin(C_2)}{L}$$

where U is the vehicle speed, L is the stopping distance, D is the wheel base, C_1 is the angle of pitch inducing bump faces and C_2 is the angle of upcoming hill faces that must be tracked by tilting the cameras. The stopping distance L, for a constant deceleration on a surface with coefficient of friction, μ, is

$$L = \frac{U^2}{2\mu g} + (t_a + 2t_c)U + B$$

where U is the initial speed, t_a is the actuation time, t_c is the compute time, and B is a buffer zone [5]. For the Demo III parameters of, U=5 m/s, μ=0.6, t_a=0.1sec, t_c=0.2 sec and B=1.0m the stopping distance is L= 5.5m.

In the situation when the vehicle is driving forward, on a bumpy road and tilting the cameras, all of these blurring effects add together giving an exposure time restriction of

$$\tau < \frac{\frac{HW}{NL^2}}{2\left(\frac{HU}{L^2} + \frac{U\sin(C_1)}{D} + \frac{U\sin(C_2)}{L}\right)}$$

If we take both C_1 and C_2 to be 60 degrees, H=1m, W=0.6m, N=10pixels, D=2m, and plot this expression we get figure 6; which shows that exposure times of less than a millisecond are necessary to meet the Demo III driving speeds of 5m/s.

These short exposure times are attainable by cooled night vision cameras, such as those based on InSb or MCT; however, all COTS uncooled technologies are unable to meet this constraint.

5.0 Image Timing Analysis

Proper image capture on a moving XUV is critical to accurate stereo matching and involves three main issues: image exposure, focal plane electronics and camera synchronization.

First, good image exposure must be produced by controlling the exposure time and aperture opening, for without it, stereo matching will be poor. The exposure time must be set short enough to prevent blur due to camera and vehicle motion and the aperture opening must be dynamically adjusted for variations is scene brightness.

Second, the focal plane electronics must provide "snapshot" image acquisition. Unfortunately, many COTS infrared cameras expose the focal plane in a rolling manner. For example, many InSb FLIR cameras expose and buffer several rows at a time, taking advantage on their fast exposure time, until an entire image is formed. This type of buffering scheme is undesirable on a moving platform since it allows as much a 1/60 of a second to elapse between exposure of the top and bottom rows in an image.

Third, it is essential for accurate stereo matching that night vision cameras used on XUVs have the capacity to accept external synchronization signals; however, most COTS infrared cameras are designed as hand held devices and lack this capacity. Without this ability, a pair of cameras can not be commanded to take simultaneous exposures producing unsynchronized images resulting in unacceptable stereo image misalignment of many pixels.

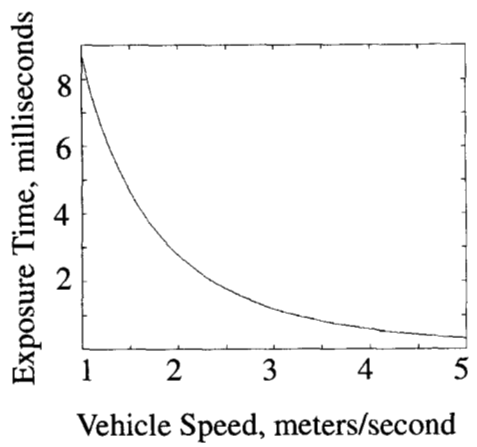


Figure 6: Necessary camera exposure time to capture clear images on a moving XUV.

6.0 Summary and Conclusion

This paper has studied the suitability of commercial off the shelf (COTS) night vision sensors for use as nighttime stereo vision sensors on board the experimen-

tal unmanned ground vehicle (XUV) being developed for the Demo III program. Nighttime field images were collected and transformed into false stereo pairs to compare the resulting disparity estimation errors and signal to noise ratios with derived requirements. A motion blur/exposure time analysis was conducted, producing a relationship between exposure time and vehicle velocity which must be respected in order to capture clear images from moving XUVs. Finally, errors due to camera synchronization were discussed.

The stereo disparity error performance analysis has shown that Gen III image intensifiers have unacceptably high noise levels for use as nighttime stereo vision cameras on board moving XUVs. Analysis of signal to noise ratios has shown that uncooled infrared sensors lack the necessary sensitivity to produce viable images for nighttime stereo vision during thermal transition periods. The motion blur analysis has shown camera exposure times must be restricted to a few milliseconds in order to capture clear images of ditches hazardous to XUV navigation. These restrictions preclude the use of present uncooled infrared and image intensifying sensors for performing nighttime stereo vision on a moving XUV. Only cooled COTS night vision sensors have the necessary sensitivity, exposure speed and synchronization capability to be used successfully as nighttime stereo vision cameras on board moving XUVs.

Acknowledgments

This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, sponsored by the Joint Robotics Program of the Office of the Secretary of Defense and managed by the Army Research Laboratory. We would like to thank Mike Grenn of NVESD for the MCT MilCam data and Sally Bennet, also of NVESD, for the loan of uncooled FLIRs for our field tests.

References

[1] L. Matthies, A. Kelly, T. Litwin. "Obstacle Detection for Unmanned Ground Vehicles: A Progress Report". International Symposium of Robotics Research, Munich, Germany October 1995.

[2] M. Hebert, R. Bolles, B. Gothard, L. Matthies, and M. Rosenblum, "Mobility for unmanned ground vehicles", in Reconnaissance, Surveillance and Target Acquisition for the Unmanned Ground Vehicle: Providing Surveillance 'Eyes' for an Autonomous Vehicle, edited by O. Firschein and T. Strat, Morgan Kaufmann Publishers, pp.95-108 1996.

[3] C. Webb and M. Norton, "Measurement Review of Infrared Systems", NVESD technical report, AMSEL:-RD-NV-TISD, 1996.

[4] M. Norton, M. Grenn, C. Webb and R. Kindsfather, "System Measurement and Performance Review of LWIR Staring Medium Format Imaging Systems", NVESD technical report, AMSEL:-RD-NV-TISD, 1997.

[5] L. Matthies and P. Grandjean, "Stochastic Performance Modeling and Evaluation of Obstacle Detectability with Imaging Range Sensors", IEEE Transactions on Robotics and Automation, Vol. 10, No. 6, Dec 1994.

[6] W. Wolf and G. Zissis, editors, "The Infrared Handbook", Environmental Research Institute of Michigan, 1978.

[7] B. Horn, "Robot Vision", MIT Press, 1986.

[8] D. Simonett and J. Estes, editors, "Manual of Remote Sensing", American Society of Photogrammetry, 1983.

[9] R. Simon, "RCA Electro-Optics Handbook", Technical Series EOH-11, 1974.

[10] L. Matthies, T. Litwin, K. Owens, A. Rankin, K. Murphy, D. Coombs, J. Gilsinn, T. Hong, S. Legowik, M. Nashman, B. Yoshimi, "Performance Evaluation of UGV Obstacle Detection with CCD/FLIR Stereo Vision and LADAR", IEEE Workshop on Perception for Mobile Agents, Santa Barbara, CA June 1998.